0570
011

#1 2 OIPE

RAW SEQUENCE LISTING   DATE: 01/02/2002
PATENT APPLICATION: US/10/014,269   TIME: 14:45:24

Input Set : **A:\um06645.txt**
Output Set: **N:\CRF3\01022002\J014269.raw**

ENTERED

```
   3 <110> APPLICANT: Nunez, Gabriel
   4        Inohara, Naohiro
   5        Ogur, Yasunori
   7 <120> TITLE OF INVENTION: NOD2 Nucleic Acids and Proteins
   9 <130> FILE REFERENCE: UM-06645
C--> 11 <140> CURRENT APPLICATION NUMBER: US/10/014,269
C--> 11 <141> CURRENT FILING DATE: 2001-10-26
  11 <160> NUMBER OF SEQ ID NOS: 52
  13 <170> SOFTWARE: PatentIn version 3.1
  15 <210> SEQ ID NO: 1
  16 <211> LENGTH: 4485
  17 <212> TYPE: DNA
  18 <213> ORGANISM: Homo sapiens
  20 <400> SEQUENCE: 1
  21 gtagacagat ccaggctcac cagtcctgtg ccactgggct tttggcgttc tgcacaaggc     60
  23 ctacccgcag atgccatgcc tgctccccca gcctaatggg ctttgatggg ggaagagggt    120
  25 ggttcagcct ctcacgatga ggaggaaaga gcaagtgtcc tcctcggaca ttctccgggt    180
  27 tgtgaaatgt gctcgcagga ggcttttcag gcacagagga gccagctggt cgagctgctg    240
  29 gtctcagggt ccctggaagg cttcgagagt gtcctggact ggctgctgtc ctgggaggtc    300
  31 ctctcctggg aggactacga gggcttccac ctcctgggcc agcctctctc ccacttggcc    360
  33 aggcgccttc tggacaccgt ctggaataag ggtacttggg cctgtcagaa gctcatcgcg    420
  35 gctgcccaag aagcccaggc cgacagccag tcccccaagc tgcatggctg ctgggacccc    480
  37 cactcgctcc acccagcccg agacctgcag agtcaccggc cagccattgt caggaggctc    540
  39 cacagccatg tggagaacat gctggacctg gcatgggagc ggggtttcgt cagccagtat    600
  41 gaatgtgatg aaatcaggtt gccgatcttc acaccgtccc agagggcaag aaggctgctt    660
  43 gatcttgcca cggtgaaagc gaatggattg gctgccttcc ttctacaaca tgttcaggaa    720
  45 ttaccagtcc cattggccct gcctttggaa gctgccacat gcaagaagta tatggccaag    780
  47 ctgaggacca cggtgtctgc tcagtctcgc ttcctcagta cctatgatgg agcagagacg    840
  49 ctctgcctgg aggacatata cacagagaat gtcctggagg tctgggcaga tgtgggcatg    900
  51 gctggacccc cgcagaagag cccagccacc ctgggcctgg aggagctctt cagcacccct    960
  53 ggccacctca atgacgatgc ggacactgtg ctggtggtgg gtgaggcggg cagtggcaag   1020
  55 agcacgctcc tgcagcggct gcacttgctg tgggctgcag ggcaagactt ccaggaattt   1080
  57 ctctttgtct tcccattcag ctgccggcag ctgcagtgca tggccaaacc actctctgtg   1140
  59 cggactctac tctttgagca ctgctgttgg cctgatgttg gtcaagaaga catcttccag   1200
  61 ttactccttg accaccctga ccgtgtcctg ttaacctttg atggctttga cgagttcaag   1260
  63 ttcaggttca cggatcgtga acgccactgc tccccgaccg accccacctc tgtccagacc   1320
  65 ctgctcttca accttctgca gggcaacctg ctgaagaatg cccgcaaggt ggtgaccagc   1380
  67 cgtccggccg ctgtgtcggc gttcctcagg aagtacatcc gcaccgagtt caacctcaag   1440
  69 ggcttctctg aacagggcat cgagctgtac ctgaggaagc gccatcatga gcccgggtg    1500
  71 gcggaccgcc tcatccgcct gctccaagag acctcagccc tgcacggttt gtgccacctg   1560
  73 cctgtcttct catggatggt gtccaaatgc caccaggaac tgttgctgca ggagggggg    1620
  75 tcccccaaga ccactacaga tatgtacctg ctgattctgc agcattttct gctgcatgcc   1680
  77 acccccccag actcagcttc ccaaggtctg ggacccagtc ttcttcgggg ccgcctcccc   1740
  79 accctcctgc acctgggcag actggctctg tggggcctgg gcatgtgctg ctacgtgttc   1800
  81 tcagcccagc agctccaggc agcacaggtc agccctgatg acatttctct tggcttcctg   1860
  83 gtgcgtgcca aaggtgtcgt gccagggagt acggcgcccc tggaattcct tcacatcact   1920
  85 ttccagtgct ctttgccgc gttctacctg gcactcagtg ctgatgtgcc accagctttg    1980
```

```
 87 ctcagacacc tcttcaattg tggcaggcca ggcaactcac caatggccag gctcctgccc   2040
 89 acgatgtgca tccaggcctc ggagggaaag gacagcagcg tggcagcttt gctgcagaag   2100
 91 gccgagccgc acaaccttca gatcacagca gccttcctgg cagggctgtt gtcccgggag   2160
 93 cactggggcc tgctggctga gtgccagaca tctgagaagg ccctgctccg gcgccaggcc   2220
 95 tgtgcccgct ggtgtctggc ccgcagcctc cgcaagcact tccactccat cccgccagct   2280
 97 gcaccgggtg aggccaagag cgtgcatgcc atgcccgggt tcatctggct catccggagc   2340
 99 ctgtacgaga tgcaggagga gcggctggct cggaaggctg cacgtggcct gaatgttggg   2400
101 cacctcaagt tgacattttg cagtgtgggc cccactgagt gtgctgccct ggcctttgtg   2460
103 ctgcagcacc tccggcggcc cgtggccctg cagctggact acaactctgt gggtgacatt   2520
105 ggcgtggagc agctgctgcc ttgccttggt gtctgcaagg ctctgtattt gcgcgataac   2580
107 aatatctcag accgaggcat ctgcaagctc attgaatgtg ctcttcactg cgagcaattg   2640
109 cagaagttag ctctattcaa caacaaattg actgacggct gtgcacactc catggctaag   2700
111 ctccttgcat gcaggcagaa cttcttggca ttgaggctgg ggaataacta catcactgcc   2760
113 gcgggagccc aagtgctggc cgaggggctc cgaggcaaca cctccttgca gttcctggga   2820
115 ttctggggca acagagtggg tgacgagggg gcccaggccc tggctgaagc cttgggtgat   2880
117 caccagagct tgaggtggct cagcctggtg gggaacaaca ttggcagtgt gggtgcccaa   2940
119 gccttggcac tgatgctggc aaagaacgtc atgctagaag aactctgcct ggaggagaac   3000
121 catctccagg atgaaggtgt atgttctctc gcagaaggac tgaagaaaaa ttcaagtttg   3060
123 aaaatcctga agttgtccaa taactgcatc acctacctag gggcagaagc cctcctgcag   3120
125 gcccttgaaa ggaatgacac catcctggaa gtctggctcc gagggaacac tttctctcta   3180
127 gaggaggttg acaagctcgg ctgcagggac accagactct tgctttgaag tctccgggag   3240
129 gatgttcgtc tcagtttgtt tgtgagcagg ctgtgagttt gggccccaga ggctgggtga   3300
131 catgtgttgg cagcctcttc aaaatgagcc ctgtcctgcc taaggctgaa cttgttttct   3360
133 gggaacacca taggtcacct ttattctggc agaggaggga gcatcagtgc cctccaggat   3420
135 agactttttcc caagcctact tttgccattg acttcttccc aagattcaat cccaggatgt   3480
137 acaaggacag cccctcctcc atagtatggg actggcctct gctgatcctc ccaggcttcc   3540
139 gtgtgggtca gtgggggccca tggatgtgct tgttaactga gtgccttttg gtggagaggc   3600
141 ccggcctctc acaaaagacc ccttaccact gctctgatga agaggagtac acagaacaca   3660
143 taattcagga agcagctttc cccatgtctc gactcatcca tccaggccat tccccgtctc   3720
145 tggttcctcc cctcctcctg gactcctgca cacgctcctt cctctgaggc tgaaattcag   3780
147 aatattagtg acctcagctt tgatatttca cttacagcac ccccaaccct ggcacccagg   3840
149 gtgggaaggg ctacacctta gcctgccctc ctttccggtg tttaagacat ttttggaagg   3900
151 ggacacgtga cagccgtttg ttccccaaga cattctaggt ttgcaagaaa aatatgacca   3960
153 cactccagct gggatcacat gtggactttt atttccagtg aaatcagtta ctcttcagtt   4020
155 aagcctttgg aaacagctcg actttaaaaa gctccaaatg cagctttaaa aaattaatct   4080
157 gggccagaat ttcaaacggc ctcactaggc ttctggttga tgcctgtgaa ctgaactctg   4140
159 acaacagact tctgaaatag acccacaaga ggcagttcca tttcatttgt gccagaatgc   4200
161 tttaggatgt acagttatgg attgaaagtt tacaggaaaa aaaattaggc cgttccttca   4260
163 aagcaaatgt cttcctggat tattcaaaat gatgtatgtt gaagcctttg taaattgtca   4320
165 gatgctgtgc aaatgttatt attttaaaca ttatgatgtg tgaaaactgg ttaatattta   4380
167 taggtcactt tgttttactg tcttaagttt atactcttat agacaacatg gccgtgaact   4440
169 ttatgctgta aataatcaga ggggaataaa ctgttgagtc aaaac                    4485
172 <210> SEQ ID NO: 2
173 <211> LENGTH: 1040
174 <212> TYPE: PRT
175 <213> ORGANISM: Homo sapiens
177 <400> SEQUENCE: 2
179 Met Gly Glu Glu Gly Gly Ser Ala Ser His Asp Glu Glu Glu Arg Ala
180 1               5                   10                  15
```

```
183 Ser Val Leu Leu Gly His Ser Pro Gly Cys Glu Met Cys Ser Gln Glu
184         20              25              30
187 Ala Phe Gln Ala Gln Arg Ser Gln Leu Val Glu Leu Leu Val Ser Gly
188         35              40              45
191 Ser Leu Glu Gly Phe Glu Ser Val Leu Asp Trp Leu Leu Ser Trp Glu
192     50              55              60
195 Val Leu Ser Trp Glu Asp Tyr Glu Gly Phe His Leu Leu Gly Gln Pro
196 65              70              75              80
199 Leu Ser His Leu Ala Arg Arg Leu Leu Asp Thr Val Trp Asn Lys Gly
200             85              90              95
203 Thr Trp Ala Cys Gln Lys Leu Ile Ala Ala Ala Gln Glu Ala Gln Ala
204         100             105             110
207 Asp Ser Gln Ser Pro Lys Leu His Gly Cys Trp Asp Pro His Ser Leu
208         115             120             125
211 His Pro Ala Arg Asp Leu Gln Ser His Arg Pro Ala Ile Val Arg Arg
212     130             135             140
215 Leu His Ser His Val Glu Asn Met Leu Asp Leu Ala Trp Glu Arg Gly
216 145             150             155             160
219 Phe Val Ser Gln Tyr Glu Cys Asp Glu Ile Arg Leu Pro Ile Phe Thr
220             165             170             175
223 Pro Ser Gln Arg Ala Arg Arg Leu Leu Asp Leu Ala Thr Val Lys Ala
224         180             185             190
227 Asn Gly Leu Ala Ala Phe Leu Leu Gln His Val Gln Glu Leu Pro Val
228         195             200             205
231 Pro Leu Ala Leu Pro Leu Glu Ala Ala Thr Cys Lys Lys Tyr Met Ala
232     210             215             220
235 Lys Leu Arg Thr Thr Val Ser Ala Gln Ser Arg Phe Leu Ser Thr Tyr
236 225             230             235             240
239 Asp Gly Ala Glu Thr Leu Cys Leu Glu Asp Ile Tyr Thr Glu Asn Val
240             245             250             255
243 Leu Glu Val Trp Ala Asp Val Gly Met Ala Gly Pro Pro Gln Lys Ser
244             260             265             270
247 Pro Ala Thr Leu Gly Leu Glu Glu Leu Phe Ser Thr Pro Gly His Leu
248     275             280             285
251 Asn Asp Asp Ala Asp Thr Val Leu Val Val Gly Glu Ala Gly Ser Gly
252     290             295             300
255 Lys Ser Thr Leu Leu Gln Arg Leu His Leu Leu Trp Ala Ala Gly Gln
256 305             310             315             320
259 Asp Phe Gln Glu Phe Leu Phe Val Phe Pro Phe Ser Cys Arg Gln Leu
260             325             330             335
263 Gln Cys Met Ala Lys Pro Leu Ser Val Arg Thr Leu Leu Phe Glu His
264         340             345             350
267 Cys Cys Trp Pro Asp Val Gly Gln Glu Asp Ile Phe Gln Leu Leu Leu
268         355             360             365
271 Asp His Pro Asp Arg Val Leu Leu Thr Phe Asp Gly Phe Asp Glu Phe
272     370             375             380
275 Lys Phe Arg Phe Thr Asp Arg Glu Arg His Cys Ser Pro Thr Asp Pro
276 385             390             395             400
279 Thr Ser Val Gln Thr Leu Leu Phe Asn Leu Leu Gln Gly Asn Leu Leu
```

RAW SEQUENCE LISTING   DATE: 01/02/2002
PATENT APPLICATION: US/10/014,269   TIME: 14:45:24

Input Set : A:\um06645.txt
Output Set: N:\CRF3\01022002\J014269.raw

```
280                    405                    410          .        415
283 Lys Asn Ala Arg Lys Val Val Thr Ser Arg Pro Ala Ala Val Ser Ala
284             420                425                430
287 Phe Leu Arg Lys Tyr Ile Arg Thr Glu Phe Asn Leu Lys Gly Phe Ser
288             435                440                445
291 Glu Gln Gly Ile Glu Leu Tyr Leu Arg Lys Arg His His Glu Pro Gly
292     450                455                460
295 Val Ala Asp Arg Leu Ile Arg Leu Leu Gln Glu Thr Ser Ala Leu His
296 465                470                475                480
299 Gly Leu Cys His Leu Pro Val Phe Ser Trp Met Val Ser Lys Cys His
300             485                490                495
303 Gln Glu Leu Leu Leu Gln Glu Gly Gly Ser Pro Lys Thr Thr Thr Asp
304             500                505                510
307 Met Tyr Leu Leu Ile Leu Gln His Phe Leu Leu His Ala Thr Pro Pro
308             515                520                525
311 Asp Ser Ala Ser Gln Gly Leu Gly Pro Ser Leu Leu Arg Gly Arg Leu
312     530                535                540
315 Pro Thr Leu Leu His Leu Gly Arg Leu Ala Leu Trp Gly Leu Gly Met
316 545                550                555                560
319 Cys Cys Tyr Val Phe Ser Ala Gln Gln Leu Gln Ala Ala Gln Val Ser
320             565                570                575
323 Pro Asp Asp Ile Ser Leu Gly Phe Leu Val Arg Ala Lys Gly Val Val
324             580                585                590
327 Pro Gly Ser Thr Ala Pro Leu Glu Phe Leu His Ile Thr Phe Gln Cys
328     595                600                605
331 Phe Phe Ala Ala Phe Tyr Leu Ala Leu Ser Ala Asp Val Pro Pro Ala
332     610                615                620
335 Leu Leu Arg His Leu Phe Asn Cys Gly Arg Pro Gly Asn Ser Pro Met
336 625                630                635                640
339 Ala Arg Leu Leu Pro Thr Met Cys Ile Gln Ala Ser Glu Gly Lys Asp
340             645                650                655
343 Ser Ser Val Ala Ala Leu Leu Gln Lys Ala Glu Pro His Asn Leu Gln
344             660                665                670
347 Ile Thr Ala Ala Phe Leu Ala Gly Leu Leu Ser Arg Glu His Trp Gly
348     675                680                685
351 Leu Leu Ala Glu Cys Gln Thr Ser Glu Lys Ala Leu Leu Arg Arg Gln
352     690                695                700
355 Ala Cys Ala Arg Trp Cys Leu Ala Arg Ser Leu Arg Lys His Phe His
356 705                710                715                720
359 Ser Ile Pro Pro Ala Ala Pro Gly Glu Ala Lys Ser Val His Ala Met
360             725                730                735
363 Pro Gly Phe Ile Trp Leu Ile Arg Ser Leu Tyr Glu Met Gln Glu Glu
364             740                745                750
367 Arg Leu Ala Arg Lys Ala Ala Arg Gly Leu Asn Val Gly His Leu Lys
368     755                760                765
371 Leu Thr Phe Cys Ser Val Gly Pro Thr Glu Cys Ala Ala Leu Ala Phe
372     770                775                780
375 Val Leu Gln His Leu Arg Arg Pro Val Ala Leu Gln Leu Asp Tyr Asn
376 785                790                795                800
```

RAW SEQUENCE LISTING   DATE: 01/02/2002
PATENT APPLICATION: US/10/014,269   TIME: 14:45:24

Input Set : A:\um06645.txt
Output Set: N:\CRF3\01022002\J014269.raw

```
379 Ser Val Gly Asp Ile Gly Val Glu Gln Leu Leu Pro Cys Leu Gly Val
380             805             810         .   815
383 Cys Lys Ala Leu Tyr Leu Arg Asp Asn Asn Ile Ser Asp Arg Gly Ile
384             820             825             830
387 Cys Lys Leu Ile Glu Cys Ala Leu His Cys Glu Gln Leu Gln Lys Leu
388         835             840             845
391 Ala Leu Phe Asn Asn Lys Leu Thr Asp Gly Cys Ala His Ser Met Ala
392     850             855             860
395 Lys Leu Leu Ala Cys Arg Gln Asn Phe Leu Ala Leu Arg Leu Gly Asn
396 865             870             875             880
399 Asn Tyr Ile Thr Ala Ala Gly Ala Gln Val Leu Ala Glu Gly Leu Arg
400             885             890             895
403 Gly Asn Thr Ser Leu Gln Phe Leu Gly Phe Trp Gly Asn Arg Val Gly
404             900             905             910
407 Asp Glu Gly Ala Gln Ala Leu Ala Glu Ala Leu Gly Asp His Gln Ser
408         915             920             925
411 Leu Arg Trp Leu Ser Leu Val Gly Asn Asn Ile Gly Ser Val Gly Ala
412     930             935             940
415 Gln Ala Leu Ala Leu Met Leu Ala Lys Asn Val Met Leu Glu Glu Leu
416 945             950             955             960
419 Cys Leu Glu Glu Asn His Leu Gln Asp Glu Gly Val Cys Ser Leu Ala
420             965             970             975
423 Glu Gly Leu Lys Lys Asn Ser Ser Leu Lys Ile Leu Lys Leu Ser Asn
424             980             985             990
427 Asn Cys Ile Thr Tyr Leu Gly Ala  Glu Ala Leu Leu Gln  Ala Leu Glu
428             995             1000            1005
431 Arg Asn  Asp Thr Ile Leu Glu  Val Trp Leu Arg Gly  Asn Thr Phe
432     1010            1015            1020
435 Ser Leu  Glu Glu Val Asp Lys  Leu Gly Cys Arg Asp  Thr Arg Leu
436     1025            1030            1035
439 Leu Leu
440     1040
443 <210> SEQ ID NO: 3
444 <211> LENGTH: 1013
445 <212> TYPE: PRT
446 <213> ORGANISM: Homo sapiens
448 <400> SEQUENCE: 3
450 Met Cys Ser Gln Glu Ala Phe Gln Ala Gln Arg Ser Gln Leu Val Glu
451 1               5               10              15
454 Leu Leu Val Ser Gly Ser Leu Glu Gly Phe Glu Ser Val Leu Asp Trp
455             20              25              30
458 Leu Leu Ser Trp Glu Val Leu Ser Trp Glu Asp Tyr Glu Gly Phe His
459         35              40              45
462 Leu Leu Gly Gln Pro Leu Ser His Leu Ala Arg Arg Leu Leu Asp Thr
463     50              55              60
466 Val Trp Asn Lys Gly Thr Trp Ala Cys Gln Lys Leu Ile Ala Ala Ala
467 65              70              75              80
470 Gln Glu Ala Gln Ala Asp Ser Gln Ser Pro Lys Leu His Gly Cys Trp
471             85              90 ·            95
```

VERIFICATION SUMMARY   DATE: 01/02/2002
PATENT APPLICATION: US/10/014,269   TIME: 14:45:25

    Input Set : **A:\um06645.txt**
    Output Set: **N:\CRF3\01022002\J014269.raw**

L:11 M:270 C: Current Application Number differs, Replaced Current Application No
L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date